1 a 9/9
b 7/9
c 4/9

2 a 7/9
b 4/9
c 9/9

3 7/9

4 a 9/9
b 3/9
c 9/9

5 8/10

Total 75/100

HW05: due Friday, December 18th

Kayla Frisoli

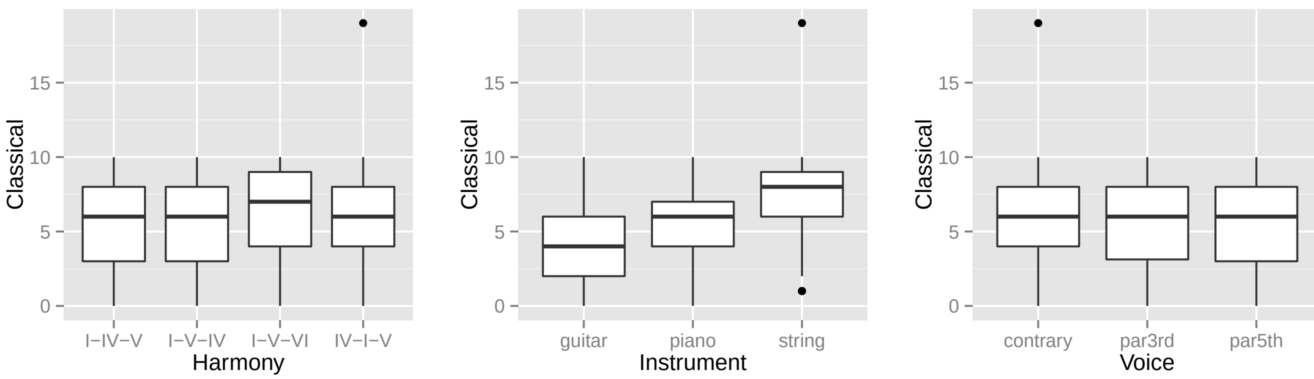
December 18, 2015

1.

(a) Influence of Instrument, Harmony & Voice on Classical ratings- linear models

Before we perform any analysis we need to explore the data.

this is great, obviously important to look at the data first...



There does not seem to be a very clear linear trend. But, we see three outliers, that must have been erroneously recorded because they are greater than the maximum value of 10. We will remove the rows containing these values. We find it is only row 1978 that must be removed.

We need to control for other variables that may have an influence on how classical a stimulus sounds.

We choose to focus on the following:

Selfdeclare - Are you a musician?
PachListen - How familiar are you with Pachelbel's Canon in D? (0 = not at all)
ClsListen - How much do you listen to classical music?
NoClass - How many music classes have you taken?

We will now fit a smaller linear model with variables we want to control for and will compare it to a model containing those variables plus the variable of interest.

Harmony

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
Selfdeclare + PachListen + ClsListen + NoClass	2175	14915.34				
Selfdeclare + PachListen + ClsListen + NoClass + Harmony	2172	14654.51	3	260.84	12.89	0.0000

Holding other variables constant, we see that harmony is a significant predictor of classical ratings F=12.89 (Pr > F = 0.0000).

Specific effects of Harmony

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.871104	0.276805	17.597603	0.000000
Selfdeclare	-0.318337	0.055083	-5.779228	0.000000
PachListen	0.200505	0.054075	3.707898	0.000214
ClsListen	0.233876	0.038119	6.135368	0.000000
NoClass	0.087779	0.039872	2.201524	0.027804
HarmonyI-V-IV	-0.061670	0.157280	-0.392105	0.695019
HarmonyI-V-VI	0.770765	0.157353	4.898331	0.000001
HarmonyIV-I-V	-0.019045	0.157280	-0.121092	0.903629

Holding all else constant, we expect, on average, that being a Harmony I-V-IV piece, compared to I-IV-V, will decrease classical rating by 0.06167.

Holding all else constant, we expect, on average, that being a Harmony I-V-VI piece, compared to I-IV-V, will increase classical rating by 0.77076.

Holding all else constant, we expect, on average, that being a Harmony IV-I-V piece, compared to I-IV-V, will decrease classical rating by 0.01905.

Instrument

We also add the following variable to this model as a covariate:

ConsInstr - How much did you concentrate on the instrument while listening?

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
Selfdeclare + PachListen + ClsListen + NoClass + ConsInstr	2174	14913.52				
Add Instrument	2172	11050.70	2	3862.82	379.62	0.0000

Holding other variables constant, we see that instrument is a significant predictor of classical ratings F=379.62 (Pr > F = 0.0000).

Specific effects of Instrument

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.533436	0.252615	13.987461	0.000000
Selfdeclare	-0.313967	0.048264	-6.505254	0.000000
PachListen	0.199726	0.046968	4.252370	0.000022
ClsListen	0.230826	0.033241	6.944066	0.000000
NoClass	0.088552	0.034645	2.555979	0.010656
ConsInstr	-0.018422	0.032738	-0.562722	0.573682
Instrumentpiano	1.442503	0.118520	12.170950	0.000000
Instrumentstring	3.244392	0.117984	27.498646	0.000000

Holding all else constant, we expect, on average, that a piano piece, compared to guitar, will be rated 1.44250 points higher in terms of classical-ness.

Holding all else constant, we expect, on average, that a string piece, compared to guitar, will be rated 3.24439 points higher.

Voice

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
Selfdeclare + PachListen + ClsListen + NoClass	2175	14915.34				
Selfdeclare + PachListen + ClsListen + NoClass + Voice	2173	14853.70	2	61.64	4.51	0.0111

Holding other variables constant, we see that voice is a significant predictor of classical ratings F=4.51 (Pr > F = 0.0111) at the .05 level, but these results are not as significant as the other variables.

Specific effects of Voice

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.277644	0.272937	19.336470	0.000000
Selfdeclare	-0.318245	0.055443	-5.740024	0.000000
PachListen	0.200475	0.054429	3.683241	0.000236
ClsListen	0.233689	0.038369	6.090626	0.000000
NoClass	0.087733	0.040133	2.186074	0.028917
Voicepar3rd	-0.389049	0.137226	-2.835102	0.004623
Voicepar5th	-0.312314	0.137179	-2.276696	0.022901

It is also good to see how the three experimental factors do when all three are included in the model - are there interactions, does one mask another? Etc...

Holding all else constant, we expect, on average, that a par3rd piece, compared to contrary, will be rated 0.389049 points lower in terms of classical-ness.

Holding all else constant, we expect, on average, that a par5th piece, compared to contrary, will be rated 0.312314 points lower.

(b) Repeated Measures Model

(i)

Hierarchical Linear Model - random intercept

$$y_i = \alpha_{j[i]} + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma^2)$$
$$\alpha_{j[i]} = \beta_0 + \eta_j, \qquad \eta_j \sim N(0, \tau^2)$$

$$rating_i = \alpha_{[participant][i]} + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma^2)$$
$$\alpha_{[participant][i]} = \beta_0 + \eta_j, \qquad \eta_j \sim N(0, \tau^2)$$

(ii) Random intercept needed?

Test using **lm command**

By itself the lm test for Subject as a fixed effect doesn't tell us very much. A positive result here, together with evidence that not many of the subject fixed effects are significant, would be good informal evidence of the need to treat subject as a random effect.

Without any other covariates

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
~1	2491	17419.89				
as.factor(Subject)	2422	12983.20	69	4436.70	12.00	0.0000

With other covariates

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
Selfdeclare + PachListen + ClsListen + NoClass	2175	14915.34				
Selfdeclare + PachListen + ClsListen + NoClass + factor(Subject)	2119	11721.84	56	3193.50	10.31	0.0000

Test using **lmer**

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
lm.1	2	11921.75	11933.39	-5958.88	11917.75			
lmer.intercept.only	3	11434.50	11451.96	-5714.25	11428.50	489.25	1	0.0000

With other covariates

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
lm.1C	6	10390.85	10424.98	-5189.43	10378.85			
lmer.intercept.var	7	10066.28	10106.09	-5026.14	10052.28	326.58	1	0.0000

7

Yes, the random intercept is definitely needed. We find much more significant results using both linear regression and lmer with the random intercept. This makes sense since we can capture more information this way!

iii. Re-examine using repeated-measures

Harmony

Control vars: Selfdeclare + PachListen + ClsListen + NoClass

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Control vars + (1 | Subject)	7	10066.28	10106.09	-5026.14	10052.28			
Control vars + (1 | Subject) + Harmony	10	10024.44	10081.31	-5002.22	10004.44	47.84	3	0.0000

Instrument

Control vars: Selfdeclare + PachListen + ClsListen + NoClass + ConsInstr

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Control vars + (1 | Subject)	8	10068.25	10113.75	-5026.13	10052.25			
Control vars + (1 | Subject) + Instrument	10	9225.56	9282.43	-4602.78	9205.56	846.70	2	0.0000

Voice

Control vars: Selfdeclare + PachListen + ClsListen + NoClass

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Control vars + (1 | Subject)	7	10066.28	10106.09	-5026.14	10052.28			
Control vars + (1 | Subject) + Voice	9	10058.93	10110.12	-5020.47	10040.93	11.35	2	0.0034

We see a similar pattern in results here as we did previously. Instrument is the most significant predictor, then harmony, then voice when controlling for other variables.

(c)

i.

	1a	1b	1c
Harmony AIC	10358.39	10042.54	10168.25
Instrument AIC	9743.08	9245.09	9046.69
Voice AIC	10385.83	10075.08	10228.40

We see that the results from 1b produce the lowest AIC for Harmony and Voice and 1c produces the lowest AIC for Instrument. We may be seeing greater changes in the AIC for Instrument because we saw from our EDA plots that instrument had the clearest relationship.

ii.

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
lmer.NOharmony	7	10066.28	10106.09	-5026.14	10052.28			
lmer.big	9	8876.85	8928.03	-4429.42	8858.85	1193.43	2	0.0000

Adding the random effects greatly reduces AIC, BIC, logLik, deviance.

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Classical ~ Selfdeclare + PachListen + ClsListen + NoClass +
##     (1 | Subject:Harmony) + (1 | Subject:Instrument) + (1 | Subject:Voice)
##    Data: ratings
##
## REML criterion at convergence: 8870
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.4821 -0.5493  0.0148  0.5162  3.4724
##
```

I'm not seeing a comparison of this model with the single-random-intercept model.

```
## Random effects:
##  Groups             Name        Variance Std.Dev.
##  Subject:Harmony    (Intercept) 0.61779  0.7860
##  Subject:Instrument (Intercept) 3.92585  1.9814
##  Subject:Voice      (Intercept) 0.05902  0.2429
##  Residual                       2.34838  1.5324
## Number of obs: 2180, groups:
## Subject:Harmony, 244; Subject:Instrument, 183; Subject:Voice, 183
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  5.04840    0.74557   6.771
## Selfdeclare -0.31897    0.15796  -2.019
## PachListen   0.19996    0.15540   1.287
## ClsListen    0.23150    0.10942   2.116
## NoClass      0.08832    0.11449   0.771
##
## Correlation of Fixed Effects:
##             (Intr) Slfdcl PchLst ClsLst
## Selfdeclare -0.154
## PachListen  -0.863 -0.212
## ClsListen   -0.042 -0.414 -0.057
## NoClass     -0.284 -0.294  0.291  0.057
```

4

The variance is highest for the instrument component, even higher than the estimated variance of the residuals. Instrument seems to be explaining much of the variation, which we expected from our previous results and exploratory graphs.

iii.

$$y_i = \alpha_1 x_i + \alpha_1 x_i + \alpha_3 x_i + \alpha_4 x_i + \alpha_{kj[i]} \qquad \epsilon_i \sim N(0, \sigma^2)$$
$$\alpha_{kj[i]} = \beta_0 + \beta_1 h_j + \beta_2 i_j + \beta_3 v_j + \eta_j, \qquad \eta_j \sim N(0, \tau^2)$$

I'm not getting this...

2.

(a)

Note: I already had added the same variables (Selfdeclare + PachListen + ClsListen + NoClass) to all models above to act as a constant/ try to account for confounding factors etc. But, here I will start with the model without any and see which variables should be added (maybe the same ones I used (quite arbitrarily) before, maybe not!).

I use the model with only a random intercept since it was best overall in terms of AIC: **Classical ~ Harmony + Instrument + Voice + (1 | Subject)**

I couldn't find this result in part 1...

```
## Data: ratings
## Models:
## lmer.1: Classical ~ Harmony + Instrument + Voice + (1 | Subject)
## lmer.2: Classical ~ Harmony + Instrument + Voice + (1 | Subject) + Popular
##        Df     AIC     BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## lmer.1 10 10434.3 10492.5 -5207.1  10414.3
## lmer.2 11  9292.3  9356.3 -4635.2   9270.3  1144      1  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7

Since we are trying to figure out what features of persons and musical passages affect classical rating, including this is not informative. It doesn't help us identify features that predict classical-ness...

Here I updated my initial model with all of the variables to check and see which ones greatly reduced AIC/BIC/Dev etc. and realized that most variables did not play a very big role. In order to keep things simple, I will only use (the very significant) popular as a covariate in the model.

lmer(Classical ~ Popular + Harmony + Instrument + Voice + (1 | Subject) , data=ratings)

(b)

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Classical ~ Popular + Harmony + Instrument + Voice + (1 | Subject)
##    Data: ratings
##
## REML criterion at convergence: 9302.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7612 -0.5833 -0.0222  0.5447  5.9290
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.779    1.334
##  Residual             2.205    1.485
## Number of obs: 2492, groups:  Subject, 70
##
## Fixed effects:
##                  Estimate Std. Error t value
## (Intercept)       8.34502    0.20850   40.02
## Popular          -0.60905    0.01595  -38.19
## HarmonyI-V-IV    -0.04777    0.08410   -0.57
## HarmonyI-V-VI     0.60511    0.08421    7.19
## HarmonyIV-I-V    -0.08258    0.08412   -0.98
## Instrumentpiano   0.79761    0.07466   10.68
## Instrumentstring  1.52981    0.08372   18.27
## Voicepar3rd      -0.29697    0.07294   -4.07
## Voicepar5th      -0.25904    0.07289   -3.55
##
## Correlation of Fixed Effects:
##             (Intr) Populr HI-V-I HI-V-V HIV-I- Instrmntp Instrmnts Vcpr3r
## Popular     -0.503
## HrmnyI-V-IV -0.203  0.005
## HrmnyI-V-VI -0.227  0.052  0.498
## HrmnyIV-I-V -0.219  0.036  0.499  0.500
## Instrumntpn -0.272  0.202  0.002  0.011  0.007
## Instrmntstr -0.401  0.497  0.001  0.025  0.018  0.524
## Voicepar3rd -0.155 -0.038 -0.002 -0.001  0.000 -0.008    -0.020
## Voicepar5th -0.155 -0.037 -0.002 -0.005 -0.003 -0.008    -0.019     0.501
```

49

I don't see a re-test of the need for the random effect here.

Look's good!

(c)

Hardly surprising but not relevant to Jiminez's questions.

9

We expect how classical a piece is to decrease with the increase of its popoularness. Being in harmony I-V-IV and IV-I-V are associated with decreased classicalness while being in harmony I-V-VI are associated with an increase. Piano and string pieces are associated with classical music as compared to the other factor here (guitar). An increase in the 3rd and 5th par (voice) are associated with a decrease in classicalness.

3.

How did you dichotomize?

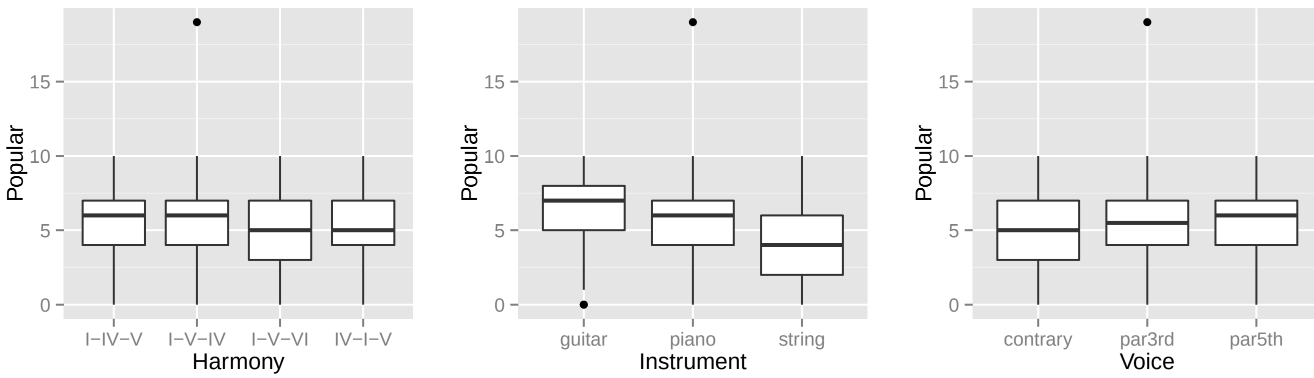
7

```
## Data: ratings
## Models:
## big.lmer: Classical ~ Popular + Harmony + Instrument + Voice + (1 | Subject)
## i2.lmer: Classical ~ Popular + Harmony:musician + Instrument + Voice +
## i2.lmer:     (1 | Subject)
##          Df    AIC    BIC  logLik deviance Chisq Chi Df Pr(>Chisq)
## big.lmer 11 9292.3 9356.3 -4635.2   9270.3
## i2.lmer  15 9272.3 9359.7 -4621.2   9242.3 27.96      4  1.271e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

After testing all possible interactions, we really only see an interaction between being a musician and harmony. The interaction plays little to no role with the other variables.

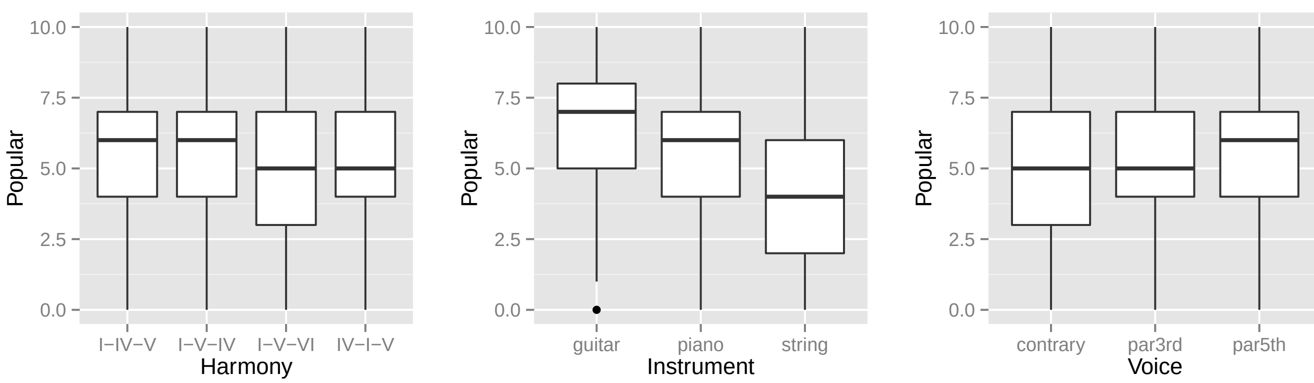
4.

The eda is helpful, thanks!



ok

There does not seem to be a very clear linear trend. But, we see three outliers (like before), that must have been erroneously recorded because they are greater than the maximum value of 10. We will remove the rows containing these values. We find it is only row 1166 that must be removed.



We at least now see a trend within the instrument (but opposite direction that we saw with classical).

9

(a)

	Estimate	Std. Error	t value
(Intercept)	6.59	0.18	36.15
HarmonyI-V-IV	-0.04	0.11	-0.40
HarmonyI-V-VI	-0.27	0.11	-2.57
HarmonyIV-I-V	-0.19	0.11	-1.79
Instrumentpiano	-0.96	0.09	-10.43
Instrumentstring	-2.61	0.09	-28.55
Voicepar3rd	0.16	0.09	1.74
Voicepar5th	0.17	0.09	1.83

We see similar (but opposite) results with popular as we did we classical in terms of the levels of the variables. But we still see that instrument is playing the biggest role with the greatest effect while harmony and voice are playing a smaller, but still influential role.

(b)

3

We expect harmony's I-V-IV, I-V-VI, and IV-I-V to be less popular than I-IV-V. We expect piano and string instrumental sounds to be less popular than guitar, and 3rd and 5th par to be more popular than contrary.

what did you find out about random effects here?

(c)

```
## Data: ratings
## Models:
## big.lmer: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
## i2.lmer: Popular ~ Harmony:musician + Instrument + Voice + (1 | Subject)
##          Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## big.lmer 10 10390 10448 -5185.1    10370
## i2.lmer  14 10376 10457 -5173.9    10348 22.494      4  0.0001598 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction significant!

```
## Data: ratings
## Models:
## big.lmer: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
## i3.lmer: Popular ~ Harmony + Instrument:musician + Voice + (1 | Subject)
##          Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## big.lmer 10 10390 10448 -5185.1    10370
## i3.lmer  13 10379 10455 -5176.7    10353 16.802      3  0.0007763 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interaction significant!

```
## Data: ratings
## Models:
## big.lmer: Popular ~ Harmony + Instrument + Voice + (1 | Subject)
## i4.lmer: Popular ~ Harmony + Instrument + Voice:musician + (1 | Subject)
##          Df   AIC   BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## big.lmer 10 10390 10448 -5185.1    10370
## i4.lmer  13 10392 10468 -5183.1    10366 4.1424      3     0.2465
```

9

Interaction not significant!

The interaction between musician and voice is not significant for either popular or classical music classifications. This may be because voice was not a very significant predictor to begin with!

5. Write-up

I have analyzed the following instrumental ratings data using conventional methods of linear models and analysis of variance but also using hierarchical methods. Hierarchical models are used to capture more information from the data at some type of group level. This can account for differences that occur within a city, or county, or state. In our situation we are looking to account for information gained within our subject, since they are each asked to listen to many pieces. You can imagine that ratings will vary from person to person. I may be more inclided to rate everything more classically than you, who tends to think everything sounds a bit more popular.

Using linear models, we found harmony, instument and voice to all be significant predictors, with different levels of the variables playing different roles. For example, guitar is more often positively associated with popular music and negatively associated with classical music.

We were able to test and see that a random intercept was needed in our model, suggesting that we needed to use some type of hierarchical model where we can account for the differences of the subjects.

We tested two different types of hierarchical models, one with a random intercept for each participant and one using a random effect of the form (1 | Subject:Instrument). The second method produced a smaller AIC for only the instrument variable, but overall the model with only the random intercept performed better, even once harmony, instrument, and voice were all included.

This seems a bit technical and sketchy for what you are trying to communicate.

There was a significant interaction between harmony and musician (binary variable for whether or not the subject considers themself a musician) for the classical model, and between harmony and musician as well as instrument and

8 musician for the popular model.

We did not include other variance components because it wouldn't be worth the loss of interpretability and we want as simple and interpretable a model as possible.

In conclusion, I found that the levels of Harmony, Instrument and Voice were extremely influential in predicting how classical or popular a stimulus sounds. The other covariates were not included since their effect didn't seem to play a large role. Instrument was the most significant and influential predictor, followed by harmony and then voice.

it looks above like some of the other covariates had significant effects - why not include them?

Classical

- We expect harmony's I-V-IV and IV-I-V to be less classical and I-V-VI to be more classical than I-IV-V
- We expect piano and string instrumental sounds to be more classical than guitar
- We expect 3rd and 5th par to be less classical than contrary

Popular

- We expect harmony's I-V-IV, I-V-VI, and IV-I-V to be less popular than I-IV-V
- We expect piano and string instrumental sounds to be less popular than guitar
- We expect 3rd and 5th par to be more popular than contrary